Express Mail B22830919

07/697830

HIGH ELECTRON MOBILITY TRANSISTOR AND METHOD OF MANUFACTURE

Background of the Invention

This invention relates generally to field effect transistors and, more particularly, to high electron mobility field effect transistors.

As is known in the art, there are several types of active devices used at microwave and millimeter wave frequencies to provide amplification of radio frequency signals. In general, one of the more common semiconductor devices used at these frequencies are field effect transistors, in particularly metal semiconductor field effect transistors (MESFETs) and high electron mobility transistors (HEMTs). Each of these transistors are provided from Group III-V materials such as gallium arsenide. What distinguishes a HEMT from a MESFET is that in a HEMT charge is transferred from a doped charge donor layer to an undoped channel layer whereas in a MESFET the charge layer and the channel layer are the same layer. Due to the presence of an undoped channel layer in a HEMT, charge transport properties of the undoped channel layer are better than those of the doped channel layer of a MESFET type structure. Accordingly, HEMTs provide higher frequency operation than MESFETs.

In a HEMT, the charge donor layer is generally a wide bandgap material, such as aluminum gallium arsenide whereas

the channel layer is a lower bandgap material, such as gallium arsenide or indium gallium arsenide. It is to be noted that bandgap refers to the potential gap between valance and conduction bands of the semiconductor materials.

In general, there are two types of HEMT structures. One type is simply referred to as a HEMT, whereas the other type is referred to as a pseudomorphic HEMT. The difference between the HEMT and the pseudomorphic HEMT is that in the pseudomorphic HEMT one or more of the layers incorporated into the HEMT structure is comprised of a material having a lattice constant which differs significantly from the lattice constants of the other materials of the device. Thus, due to resulting lattice constant mismatch, the crystal structure of the material providing the channel layer is strained.

In a HEMT structure, charge is transferred from the donor layer to an undoped channel layer. For Group III-V materials, a doped charge donor layer is comprised of a wide bandgap material, such as gallium aluminum arsenide, whereas the channel layer is typically comprised of a material having better electron transport properties. Typically, a lower bandgap material, such as gallium arsenide is used. In a pseudomorphic HEMT, the undoped gallium arsenide channel layer is replaced by a channel layer comprised of a lower bandgap material, such as indium gallium arsenide. In

either event, however, each of the HEMT and pseudomorphic HEMT structures are used to provide amplification of high frequency microwave and millimeter wave signals.

For low noise and high frequency applications of high electron mobility transistors, it is important to have a narrow recess disposed through the contact layers of the device and over the charge donor layer. That is, the recess opening is preferably only slightly longer than the gate length of the gate electrode disposed within the recess. This arrangement has provided HEMTs and pseudomorphic HEMTs that have relatively high frequency operating characteristics and relatively low noise figures. For power applications in MESFETs, it is generally known that a recessed opening larger than the gate is necessary to provide the MESFET having relatively high gate to drain breakdown voltage characteristics.

Returning to a HEMT, on the etched gallium aluminum arsenide surface which is generally the upper surface in most HEMT structures, there exists a large number of surface states. Such surface states also exist on the GaAs surface. Some authors have estimated the surface states to be as many

BH31 as 10^{14} cm^{-2}. These states most likely arise from gallium and aluminum oxides. It has been suggested that these states once occupied, increase the gate to drain breakdown voltage characteristic by capturing electrons and thus

4

decreasing the electric field concentrated at the gate metal edge on the drain side of the transistor.

The breakdown voltage characteristics of high electron mobility transistors has limited their use to relatively low power, low noise applications. This follows since the output impedance of a HEMT is generally related to the drain bias level. Low breakdown voltage characteristics limits the operating drain voltage of a HEMT. For a given DC power level, it is general advantageous to bias a HEMT for high power applications at relatively high drain voltages and low drain current rather than vice versa. Biased at high drain voltage provides a higher output impedance for the HEMT and therefore a more easy impedance match to a 50 ohm system characteristic impedance which is generally encountered in most applications. In particular, this match is more easily made over broad ranges of operating frequencies. Further, to provide high levels of RF voltage gain from such a device it is generally necessary to operate the device at a relatively high drain voltage DC bias. However, as indicated above, although it would be desirable to bias HEMTs at higher breakdown voltage, such is generally not possible since the HEMTs have relatively low breakdown voltage characteristics.

Therefore, high electron mobility transistors are used in relatively low power, low noise, applications, because

the known high electron mobility transistors generally have relatively low gate to drain reverse breakdown voltage characteristics. This situation is undesirable since the high frequency characteristics of HEMTs and the relatively high gain of HEMTs in comparison to MESFETs, would otherwise be useful for higher power applications.

CLOIC P

Summary of the Invention

In accordance with the present invention, a high electron mobility transistor includes a charge donor layer comprised of a first Group III-V material and a channel layer disposed adjacent to said charge donor layer and comprised of a second Group III-V material having a bandgap energy lower than the bandgap energy of said first Group III-V material. The high electron mobility transistor further includes a pair of ohmic contacts disposed over a first one of said charge donor and channel layers and a Schottky barrier contact disposed on a second one of said charge donor and channel layers. The high electron mobility transistor further includes means for shielding or screening the channel layer from surface charges present in the region between the gate and drain electrodes of the high electron mobility transistor. With such an arrangement, by providing a shield to screen the channel layer from surface charge, the parasitic effects associated with surface charges in the gate to drain region of the transistor are reduced thereby permitting higher breakdown voltage characteristics from the high electron mobility transistor without degrading the RF performance characteristics of the HEMT.

We believe that breakdown voltage limitations in high electron mobility transistors are at least partially related to surface states or excess surface charge present in the

7

charge donor layer or channel layer of the transistor. These surface states are the same surface states generally believed to alleviate the high electric field between the gate metal edge and the charge donor layer on the drain side of the device. Electrons can spatially tunnel from the metal gates to the nearby surface states in the semiconductor layers, such as the charged donor layer. However, as the gate to drain voltage increases, those surface states which are spatially removed from the gate electrode also become ionized. These trapped electrons thus provide a depletion region which is generally commensurate with the density of the surface states. The re-emission of these trapped charges requires a finite transit time which is generally much greater than the transit time of charge response under the gate electrode. Thus, the depletion region is essentially static relative to the input RF signal and as the RF signal increases in frequency the effects become more pronounced. Thus, the presence of the surface charges provide, in essence, a parasitic gate. The parasitic gate electrode thus acts as a slow gate which does not respond to high frequency signals as does the regular gate electrode of the HEMT. Accordingly, the presence of this gate electrode reduces the overall efficiency of the transistor. Further, we have also observed that the RF characteristics of a transistor can vary dramatically with

4

quiescent bias voltages if surface states are present. The greater the drain gate voltage the more surface states are filled and the greater the large signal degradation that is provided.

In accordance with a further aspect of the present invention, the high electron mobility transistor further includes the charge donor layer comprised of the high bandgap Group III-V material with said charge donor layer having a sheet of dopant disposed substantially close to the interface provided by the charge donor layer and the channel layer. The dopant sheet has a concentration of dopant atoms confined to a few atomic layer thicknesses of the layer of material. With this particular arrangement, by providing the dopant sheet, the dopant has been removed from the vicinity of the gate electrode which results in a reduction of avalanche breakdown effects between the gate and drain regions of the high electron mobility transistor.

In accordance with a further aspect of the present invention, the shielding means is a pair of charge screen layers comprised of relatively lightly doped materials disposed between the gate electrode and source and drain electrodes of the high electron mobility transistor. With such an arrangement, by providing charge screen layers disposed between gate and source and drain electrodes, each of the charge screen layers supply positive space charge

that compensate for the negative surface states present in the active layer.

In accordance with a still further aspect of the present invention, the high electron mobility transistor further includes a pair of charge screen layers disposed over the first one of active and charge donor layers of the high electron mobility transistor. The pair of screen layers are patterned to provide a double recessed channel. A first charge screen layer disposed adjacent to the charge donor layer is etched to provide a recess having a first length between source and drain electrodes, whereas a second charge screen layer disposed over the first aforementioned charge screen layer, as well as, a portion of the aforementioned first charge screen layer are etched to provide a second, substantially longer length between source and drain electrodes. The gate electrode is provided in the first aforementioned recess in Schottky barrier contact with the charge donor layer. With this particular arrangement, by providing a double recessed structure, the structure reduces the effect of surface states between the gate electrode and the active layer of the high electron mobility transistor by further removing such regions having such states from the immediate vicinity of the gate. This arrangement thus provides HEMTs having relatively high gate-drain breakdown voltage characteristics. Further, however,

10

the presence of the lightly doped charge screen layers permits the above device to exhibit good breakdown voltage levels at relatively high frequencies.

11

Brief Description of the Drawings

The foregoing features of this invention, as well as the invention itself, may be more fully understood from the following detailed description of the drawings, in which:

FIGs. 1A-1E are a series of cross-sectional views showing the steps in fabricating a high electron mobility transistor in accordance with the one aspect of the present invention, and

FIG. 2 is a cross-sectional view showing a high electron mobility transistor in accordance with a further aspect of the present invention.

Description of the Preferred Embodiments

Referring now to FIG. 1A, a substrate 12, here comprised of semi-insulating gallium arsenide or other suitable semiconductor material is shown having a plurality of layers disposed thereon. In particular, disposed over substrate 12 is a first layer 14, comprised of a superlattice buffer layer comprised of alternating layer pairs (not shown) of gallium arsenide and aluminum gallium arsenide, each one of said layers having a typical thickness of 50-100Å disposed to provide a superlattice as is known in the art. Disposed over superlattice 14 is a undoped layer 16 of, here, gallium arsenide. Disposed over layer 16 is a back layer 18 comprised of aluminum gallium arsenide which provides a large bandgap material on the back or bottom surface of a channel layer 20, as will be described shortly. Back layer 18 has a pair of regions 18a, 18b. Region 18a is a first region or spacer layer region of generally undoped wide bandgap material such as $Al_xGa_{1-x}As$ where X is typically between 0.1 and 0.3, whereas region 18b is a second region of undoped, wide bandgap material here also AlGaAs. Such regions 18a, 18b have disposed therebetween a region of high dopant concentration 19, here commonly referred to ~~as a~~ an n type dopant spike and which is preferably described as a substantially two-dimensional layer or sheet of dopant material having a dopant concentration in the typical range

13

of $0.3X10^{12}$ to $5X10^{12}$ a/cm^2 with the range of $0.3X10^{12}$ to $1.5X10^{12}$ being preferred. Further, the dopant layer 19 is confined to several angstroms of thickness of composite layer 18 thus to provide a sheet having a thickness of a few atomic layers. Dopant layer 19 is spaced some 30-50Å by region 18a from the interface or heterojunction between region 18a and a layer 20 which provides the channel layer for the device to be fabricated. The channel layer 20 is comprised of a lower bandgap material, such as gallium arsenide or indium gallium arsenide or other suitable materials as is known. Layer 20 is undoped and provides a region of low impurity concentration and thus provides a region where carrier mobilities are relatively high.

Disposed over channel layer 20 is a second wide bandgap material layer 22, having an undoped spacer region 22a having a typical thickness of 30Å to 50Å and a lightly doped screening region 22b which are separated by ~~a~~ an n type dopant spike region 21, having a sheet doping concentration of typically $2X10^{12}$ to $5X10^{12}$ a/cm^2 and provides the charge donor region for the channel layer 20. Region 22b is lightly doped having a dopant concentration in the range of $1X10^{17}$ to $5X10^{17}$ a/cm^3 typically and a thickness typically of 200Å to 400Å and provides a first charge screen layer, as will be further described. Disposed over region 22b is a second charge screen layer 24 here comprised of a medium bandgap

14

B33H L521

B33H 14

material such as gallium arsenide (GaAs) having a dopant concentration of $1X10^{17}$ to $5X10^{17}$ a/cm^3. Layer 24 has a thickness typically in the range of 300Å to 400Å. Of course, for both layers 22 and 24 other thicknesses could be used. Disposed over layer 24 is a contact layer 26 comprised of a relatively highly doped layer (i.e. having a dopant concentration generally greater than $1X10^{18}$ a/cm^3 or greater) of a Group III-V material such as gallium arsenide which is used to provide ohmic contact to source and drain electrodes, as will be described below.

Referring now to FIG. 1B, layer 26 has disposed thereover a patterned photoresist layer 42 which defines regions for source and drain contacts. Layer 42 is patterned to provide regions or apertures 42a, 42b, as shown, and a layer of metal is introduced into apertures 42a, 42b and over portions of photoresist 42. Thereafter, the photoresist layer 42 is removed leaving behind electrodes 28a, 28b in the regions of apertures 42a, 42b, as shown. Such electrodes provide ohmic source and drain contacts to layer 26. Preferably the electrodes 28a, 28b are alloyed with layer 26 using techniques conventionally known to provide low resistivity ohmic type contacts to layer 26 and layers there below including underlying portion of the channel layer 20.

15

Referring now to FIG. 1C, photoresist layer 42 (FIG. 1B) having been removed leaves behind source and drain contacts 28a, 28b, respectively as shown. A second photoresist layer 44 is deposited over and is patterned to provide an aperture 44a which, here, is patterned to provide a relatively wide aperture between source and drain contacts 28a, 28b, as shown. The structure 10 is brought in to contact with an etchant which will suitably etch away exposed portions of layer 26, as well as, the selective portion of underlying layer 24, as shown. This provides a first recess through layers 26 and 24, as shown, for eventual formation of a gate electrode as will be described in conjunction with FIG. 1E. A suitable etchant for layers 26 and 24 is a dilute solution of ammonia hydroxide and hydrogen peroxide which is held in contact with said layers for a selected period of time.

Referring now to FIG. 1D, photoresist layer 44 (FIG. 1C) is removed and is replaced with a third photoresist layer 46 which is likewise patterned to, here, provide an aperture 46a in photoresist layer 46 generally disposed within the recess portion 25′ of the layers 24 and 26. Aperture 46a exposes a selective underlying portion of layer 24. An etchant such as a dilute solution of ammonia hydroxide and hydrogen peroxide is brought into contact with the exposed portion of layer 24 to form a recess 23 within

the aforementioned recess portion 25′. Here recess 23 is provided through layer 24 and through a selected portion of layer 22. This mask is also used to deposit a Schottky barrier metal (not shown) to provide a gate electrode in Schottky barrier contact with layer 22 as generally shown in FIG. 1E.

The structure 10 shown in FIG. 1E includes a pair of charge donor layers 18, 22, disposed over and under channel layer 20. This is a so-called high breakdown high electron mobility transistor structure. Further, the structure described in conjunction with FIG. 1E includes a pair of charge screening layers 22b, 24 disposed between the contact layer 26 and the dopant spike or charge donor region 21. Here region 22b differs from the conventional charge donor layer of a typical high electron mobility transistor since it has a generally low dopant concentration except for the dopant spike region 21 disposed substantially adjacent the heterojunction provided between layers 20 and 22. The heterojunction is formed in said layers by disposing the spacer region of the wide bandgap material, such as aluminum gallium arsenide characteristic of spacer region 22a with the lower bandgap material, such as gallium arsenide or indium gallium arsenide used for layer 20.

With this arrangement of the layer 22, a substantial portion of the charge is removed from the vicinity of the

17

gate electrode 30 which is accomplished by the relatively low dopant concentration throughout most of the layer 22 except for the dopant spike regions 21.

The presence of the double wide recessed gate electrode in the pair of charge screen regions or layers 22b, 24 serves to further reduce the influence of surface states or a resulting surface inversion layer which exists on the surface of the semiconductor regions disposed between the gate electrode 30 and drain electrode 28b from providing parasitic gate effects in the channel layer 20. Furthermore, the lightly doped charge screen layers 22b and 24 further reduces the effect of surface states between the gate electrode and the channel layer 20. With each of the aforementioned arrangements therefore since the parasitic effects between the gate electrode and the channel layer are reduced, the resulting high electron mobility transistor can be operated at substantially higher drain bias voltage than conventionally known HEMTs.

As also shown in conjunction with FIG. 1E, a ground plane conductor 11 is disposed over an opposing surface of substrate 12, as is conventionally provided.

Referring now to FIG. 2, a high electron mobility transistor 50 having a single charge donor layer and having the structure as otherwise generally described in conjunction with FIG. 1 is shown. Here the transistor is

18

fabricated using similar approaches generally described above. Thus, here the transistor 50 includes a substrate 52 having a ground plane 51 disposed over the first surface thereof and a buffer layer 54, here a superlattice type buffer layer, disposed over a second surface of substrate 52. Disposed over the superlattice buffer layer 54 is a channel layer 56 comprised of a relatively low or narrow bandgap material such as gallium arsenide or indium gallium arsenide. Disposed over channel layer 56 is a charge donor layer 58 comprised of a dopant spike region 59 having the general characteristics as discussed above in conjunction with FIG. 1E for regions 19 and 21. That is, dopant region 59 is generally characterized as a sheet or two-dimensional layer of dopant material having a dopant concentration of B33H $2X10^{12}$ to $5X10^{12}$ a/cm^2. Dopant region 59 is spaced from B521 channel layer 56 by a spacer layer 58a (30Å to 50Å in thickness) of undoped wide bandgap material such as gallium aluminum arsenide. The remainder of layer 58, here region B521 58b (i.e. typically 200Å to 400Å), is, again, lightly doped gallium aluminum arsenide. Disposed over region 58b of layer 58 is a layer 64 comprised of lightly doped N-type gallium arsenide as for layer 24. Disposed over layer 64 is a relatively heavy doped layer of N-type gallium arsenide to provide a contact with source and drain electrodes as shown. Gate electrode 70 is disposed in a double recessed portion

of layers 66 and 64, as also shown, and as generally described above in conjunction with FIG. 1E. The structure shown in FIG. 2 is a high electron mobility transistor having a single charge donor layer. The presence of the double wide recess, the dopant spike region 59, and the pair of charge screen regions or layers 58b and 64 provide individually and collectively improvements in breakdown voltage characteristics as generally described for the structure shown in FIG. 1E.

Such structures could alternatively be used with so-called inverted high electron mobility transistors which have a charge donor layer under rather than over the channel layer as generally described above.

Having described preferred embodiments of the invention, it will now become apparent to one of skill in the art that other embodiments incorporating their concepts may be used. It is felt, therefore, that these embodiments should not be limited to disclosed embodiments, but rather should be limited only by the spirit and scope of the appended claims.